Package ‘MaximinInfer’

April 12, 2023

Type Package
Title Inference for Maximin Effects in High-Dimensional Settings
Version 2.0.0
Author Zhenyu Wang; Zijian Guo
Maintainer Zhenyu Wang <zw425@stat.rutgers.edu>
Description Implementation of the sampling and aggregation method for the covariate shift maximin effect, which was proposed in <arXiv:2011.07568>. It constructs the confidence interval for any linear combination of the high-dimensional maximin effect.
License GPL-3
Encoding UTF-8
RoxygenNote 7.2.3
Suggests knitr, rmarkdown
Imports MASS, stats, CVXR, glmnet, intervals, SIHR
Depends R (>= 2.10)
NeedsCompilation no
Repository CRAN
Date/Publication 2023-04-12 09:40:02 UTC

R topics documented:

- decide_delta .......................... 2
- Infer .................................. 2
- Maximin ................................ 3
- measure_instability ..................... 5

Index 6
decide_delta  
*Decide ridge penalty data-dependently*

**Description**

To tell if the estimator is stable or not without ridge penalty at first. If instable, it picks a ridge penalty data-dependently.

**Usage**

```r
decide_delta(
  obj,
  gen.size = 500,
  step_delta = 0.1,
  MAX_iter = 100,
  verbose = FALSE
)
```

**Arguments**

- **obj**: The returned list of Maximin
- **gen.size**: The generating sample size (Default = 500)
- **step_delta**: The step size of searching delta (Default = 0.1)
- **MAX_iter**: Maximum of iterations for searching (Default = 100)
- **verbose**: Print information about delta and reward (Default = FALSE)

**Value**

- **delta**: The data-dependent ridge penalty
- **reward.ratio**: The ratio of penalized reward over non-penalized reward

---

**Infer**  
*Inference method*

**Description**

Given the returned list of Maximin, compute the Point estimator and Confidence interval.
**Maximin**

Usage

```r
Infer(
  obj,
  delta = 0,
  gen.size = 500,
  threshold = 0,
  alpha = 0.05,
  alpha.thres = 0.01
)
```

Arguments

- `obj`: returned list of Maximin
- `delta`: The ridge penalty (Default = 0)
- `gen.size`: The generating sample size (Default = 500)
- `threshold`: Should generated samples be filtered or not? if 0, use normal threshold to filter; if 1, use chi-square threshold to filter; if 2, do not filter (Default = 0)
- `alpha`: confidence value to construct confidence interval (Default = 0.05)
- `alpha.thres`: confidence value to select generated samples (Default = 0.01)

Value

- `weight`: The weight vector for groups, of length \( L \)
- `mm.effect`: The aggregated maximin effect (coefficients), of length \( p \) or \( p + 1 \)
- `mminfer`: The list of length \( n\text{loading} \), each contains the point estimator and confidence interval

---

Maximin

*Returns a list that provides materials for later inference method.*

Description

Given list of observations, compute the bias-corrected initial estimators and do bias-correction to the regressopm covariance matrix.

Usage

```r
Maximin(
  Xlist,
  Ylist,
  loading.mat,
  X0 = NULL,
  cov.shift = TRUE,
  cov0 = NULL,
  intercept = TRUE,
)```
intercept.loading = FALSE,
lambda = NULL,
verbose = FALSE
)

Arguments

Xlist  list of design matrix for source data, of length \( L \)
Ylist  list of outcome vector for source data, of length \( L \)
loading.mat Loading matrix, of dimension \( n.loading \times p \), each column corresponds to a loading of interest
X0 design matrix for target data, of dimension \( n0 \times p \) (default = NULL)
cov.shift Covariate shifts or not between source and target data (default = TRUE)
cov0 Covariance matrix for target data, of dimension \( p \times p \) (default = NULL)
intercept Should intercept be fitted for the initial estimator (default = TRUE)
intercept.loading Should intercept term be included for the loading (default = FALSE)
lambda The tuning parameter in fitting initial model. If NULL, it will be picked by cross-validation. (default = NULL)
verbose Should intermediate message(s) be printed. (default = FALSE)

Details

The algorithm implemented scenarios with or without covariate shift. If \( \text{cov0} \) is specified, the \( X0 \) will be ignored; if not, while \( X0 \) is specified, \( \text{cov0} \) will be estimated by \( X0 \). If both are not specified, the algorithm will automatically set \( \text{cov.shift} \) as \( \text{FALSE} \).

Value

The returned list contains the following components:

- \( \text{Gamma.plugin} \): The plugin regression covariance matrix
- \( \text{Gamma.debias} \): The proposed debiased regression covariance matrix
- \( \text{Var.Gamma} \): The variance matrix for sampling the regression covariance matrix
- \( \text{fits.info} \): The list of length \( L \), that contains the initial coefficient estimators and variance of fitted residuals.
- \( \text{Points.info} \): The list of length \( L \), that contains the initial debiased estimator for linear combinations and its corresponding standard error.

Examples

\[
L = 2
n1 = n2 = 100; p = 4
X1 = MASS::mvrnorm(n1, rep(0,p), Sigma=diag(p))
X2 = MASS::mvrnorm(n2, rep(0,p), Sigma=0.5*diag(p))
b1 = seq(1,4)/10; b2 = rep(0.2, p)
\]
$$y_1 = \text{as.vector}(X_1 \%*% b_1 + \text{rnorm}(n_1)); \quad y_2 = \text{as.vector}(X_2 \%*% b_2 + \text{rnorm}(n_2))$$
loading1 = rep(0.4, p)
loading2 = c(-0.5, -0.5, rep(0, p-2))
loading.mat = cbind(loading1, loading2)
cov0 = diag(p)
mm = \text{Maximin}(\text{list}(X_1, X_2), \text{list}(y_1, y_2), \text{loading.mat}, \text{cov0}=\text{cov0})

# inference
out = \text{Infer}(mm, \text{gen.size}=10)

---

**measure_instability**  
*measurement of instability*

**Description**
compute the instability measurement given a specific ridge penalty

**Usage**

```r
measure\_instability(
  obj,
  delta = 0,
  gen.size = 500,
  threshold = 0,
  alpha.thres = 0.01
)
```

**Arguments**

- **obj**: The returned list of Maximin
- **delta**: The ridge penalty (Default = 0)
- **gen.size**: The generating sample size (Default = 500)
- **threshold**: Should generated samples be filtered or not? if 0, use normal threshold to filter; if 1, use chi-square threshold to filter; if 2, do not filter. (Default = 0)
- **alpha.thres**: The confidence value to select generated samples (Default = 0.01)

**Value**
The measurement of instability
Index

decide_delta, 2
Infer, 2
Maximin, 3
measure_instability, 5